IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re U.S. Patent Application of	)	
	)	
AIZAWA et al.	)	**Unit 2655**
	)	
Application Number: 10/782,925	)	
	)	
Filed: February 23, 2004	)	
	)	
For: DISK ARRAY SYSTEM AND A METHOD OF AVOIDING FAILURE OF THE DISK ARRAY SYSTEM	)	
	)	
ATTORNEY DOCKET NO. HITA.0522	)	

Honorable Assistant Commissioner
for Patents
Washington, D.C. 20231

COVER LETTER

Sir:

The below-identified communications are submitted in the above-captioned application or proceeding:

[x] Petition to Make Special under 37 CFR §1.102(d) for Accelerated Examination
[ ] Substitute Specification
[ ] Assignment
[ ] Terminal Disclaimer
[x] Information Disclosure Statement
[x] Statements & Pre-exam search report

[ ] Please charge my **Deposit Account Number** _____ in the amount of _____ to cover the fees for _____. A duplicate copy of this paper is enclosed.

[x] A check in the amount of **$130.00** to cover the petition fee is enclosed.

[x] The Commissioner is hereby authorized to charge any additional fees associated with this communication, or credit any overpayment to **Deposit Account Number 08-1480**.

Respectfully submitted,

Stanley P. Fisher
Registration Number 24,344

Juan Carlos A. Marquez
Registration No. 34,072

REED SMITH LLP
3110 Fairview Park Drive
Suite 1400
Falls Church, Virginia 22042
(703) 641-4200
April 18, 2005

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re U.S. Patent Application of	)	
	)	
AIZAWA et al.	)	**Unit 2655**
	)	
Application Number: 10/782,925	)	
	)	
Filed: February 23, 2004	)	
	)	
For: DISK ARRAY SYSTEM AND A METHOD OF AVOIDING FAILURE OF THE DISK ARRAY SYSTEM	)	
	)	
ATTORNEY DOCKET NO. HITA.0522	)	

Honorable Assistant Commissioner
for Patents
Washington, D.C. 20231

PETITION TO MAKE SPECIAL UNDER 37 C.F.R. § 1.102(d)

FOR ACCELERATED EXAMINATION

Sir:

Pursuant to 37 C.F.R. § 1.102(d), Applicants respectively request that the application to be examined on the merits in conjunction with the pre-examination search results, the detailed discussion of the relevance of the results and amendments as filed concurrently.

Substantive consideration of the claims is respectfully solicited. Should there be any outstanding issues requiring discussion that would further the prosecution and allowance of the above-captioned application, the Examiner is invited to contact the Applicants' undersigned representative at the address and telephone number indicated below.

Respectfully submitted,

Stanley P. Fisher
Registration Number 24,344

Juan Carlos A. Marquez
Registration Number 34,072

REED SMITH LLP
3110 Fairview Park Drive
Suite 1400
Falls Church, Virginia 22042
(703) 641-4200
April 18, 2005
SPF/JCM/JT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re U.S. Patent Application of	)	
	)	
AIZAWA et al.	)	**Unit 2655**
	)	
Application Number: 10/782,925	)	
	)	
Filed: February 23, 2004	)	
	)	
For: DISK ARRAY SYSTEM AND A METHOD OF AVOIDING FAILURE OF THE DISK ARRAY SYSTEM	)	
	)	
ATTORNEY DOCKET NO. HITA.0522	)	

Honorable Assistant Commissioner
for Patents
Washington, D.C. 20231

STATEMENTS & PRE-EXAMINATION SEARCH REPORT SUPPLEMENTAL TO THE PETITION TO MAKE SPECIAL

Sir:

Pursuant to 37 C.F.R. §§ 1.102 and MPEP 708.02 VIII, Applicants hereby submit that (1) all claims of record are directed to a single invention, or if the Office determines that all the claims presented are not obviously directed to a single invention, will make an election without traverse as a prerequisite to the grant of special status; (2) a pre-examination search has been conducted according to the following field of search; (3) copies of each reference deemed most closely related to the subject matter encompassed by the claims are enclosed; and (4) a detailed discussion of the references pointing out how the claimed subject matter is patentable over the references is also enclosed herewith.

FIELD OF THE SEARCH

The field of search covered Class 714, subclasses 1, 5, 100, 770 and 799 (U.S. & Foreign). Additionally, a computer database search was conducted on the U.S.P.T.O. systems EAST and WEST for U.S. and foreign patents; a keyword search was conducted in Class 709, Class 711, and Class 714 (U.S. & Foreign); and a literature search was also

conducted on the internet for relevant non-patent documents. Examiner Stephen Baker in Class 714 (Art Unit 2133) was consulted in confirming the field of search.

The search was directed toward a disk array system and a method of avoiding failure of the disk array system. In particular, the search was directed towards claims 1, 11 and 17 of U.S. Application Number 10/782,925. With reference to the specification, FIG. 1 illustrates a disk array system 10 including channel adapters CHA 11, shared memory 13, cache memory 14, switch unit 15, and disk adapters 12 connected to disk drives 16 (p. 16, lines 12 – 18). Disk drives 16 may be configured as a RAID (p. 18, lines 4 – 12). Service processor SVP 2 observes and manages disk array system 10 (p. 18, lines 17 – 21). FIG. 5 illustrates a schematic, wherein when a read or write error exceeds a threshold, it is judged that there may be a risk of failure in the disk drive 16 (#4) (p. 23, lines 7 – 18). The contents are then read to cache memory 14, and then to spare disk drive SP 16. *Id.* When data copy to disk drive 16 (SP) is started, an unused RAID group in RAID group 17(S) is reserved, then RAID group 17(P) (which includes SP 16) is paired with the reserved RAID group in group 17(S) (p. 23, lines 19 – 25). This pairing information is then registered in pairing information management table T2 (Figs. 3A-B; p. 24, lines 2 – 3). During the data transfer to spare disk drive 16 (SP), the host 1 may request writing (which would normally go to primary volume 18(P) of group 17(P)). However, instead, data writing is provided to secondary volume 18(S) (p. 24, lines 4 – 10). With reference to the claims, a disk array system or method comprises a first control unit (or step) observing a frequency of disk access error above a threshold and copying data into cache memory in response, a second control unit (or step) processing access requests and making backup storage during the copying, and a third control unit (or step) copying data from the backup storage.

LIST OF RELEVANT REFERENCES

The search revealed the following U.S. patents or applications, which are listed for convenience:

U.S. Patent Number	Inventor(s)
5,579,474	Kakuta et al.
6,442,711	Sasamoto et al.
6,859,888	Furuya et al.
2002/0162057	Talagala

Published Patent Application	Inventor(s)
2003/0233613	Ash et al.
2004/0153844	Ghose et al.

Discussion of References:

U.S. Patent No. 5,579,474 to **Kakuta** et al. describes a disk array system and control method wherein write data is written in duplicate fields for scheduling of writing data. FIG. 1 illustrates a CPU 1, a disk array controller (ADC) 2, and a disk array unit (ADU) 3 (col. 4, lines 18 – 29). The ADC 2 includes a channel path director 5, and a cache memory 7 storing data and an address translation table (col. 4, lines 31 – 37). When a command is issued from the CPU 1, the channel path director 5 of the ADC 2 judges whether the command can be received or not (col. 4, lines 49 – 51). FIG. 5 illustrates a write process, such that after a microprocessor (MP1) 20 of the disk array controller (ADC) 2 has reported the tentative write completion to the CPU 1, the MP1 20 monitors the read/write requests (I/O state) of the CPU 1 (Step 512) (col. 11, lines 43 – 50). The MP1 20 counts the number of read/write requests per unit time issued by the CPU 1 relative to the subject logical group 10. If this number is smaller than a preset number and if the CPU 1 does not presently issue a read/write request to the logical group of an SCSI drives 12, the operation starts for generating the parity and writing in the SCSI drive 12. **Kakuta** merely monitors the numbers of read/write requests of the CPU 1 so to decide the timing for generating a new parity in the SCSI drive 12. **Kakuta** is *determined* to generate a new parity in the SCSI drive 12 (will happen no matter what and just a matter of timing), rather than backing-up a disk drive or generating a new parity *only if* a frequency of disk access error reaches a threshold (a non-timing condition which will not necessarily happen). As such, **Kakuta** does not provide "a first control unit provided in the disk adapter for observing occurrence of access error with respect to the data disk drives, the first control unit, when the frequency of occurrence of the access error exceeds a predetermined threshold, copying data stored in the data disk drive exceeding the threshold in the spare disk drive via the cache memory," "a second control unit provided in the disk adapter for processing access request directed to the RAID group during the copying process by the first control unit, the second control unit making the backup storage take over a write request directed to the RAID group," and "a third control unit provided in the disk adapter for copying data written in the backup storage by the second control unit to the data disk drives

and the spare disk drive other than the data disk drive exceeding the threshold when the copying process by the fist control unit is finished" as recited in claim 1 and paraphrased as corresponding steps in claims 11 and 17.

U.S. Patent No. 6,442,711 to **Sasamoto** et al. describes a system and method for avoiding storage failures in a disk array system. FIG. 1 illustrates a host computer 100 connected to a RAID based disk array control unit 101, which includes a protocol controller 111 (col. 3, lines 8 – 15). A disk array 102 has a plurality of disk drives 103 and stores user data in a RAID manner (col. 3, lines 15 – 23). The disk array control unit 101 includes a buffer controller 110 that: (i) manages an Input/Output (I/O) operation of a buffer memory 109, (ii) creates parity data, and (iii) regenerates lost user data from the parity data (col. 3, lines 36 – 48). FIG. 5 illustrates a disk maintenance procedure wherein a calculating means 302 calculates an error rate and a judging means 303 judges if the disk drive needs to be changed (col. 5, lines 36 – 60). If a calculated values exceeds a threshold, an executing means 304 regards the disk drive as half-failed (step 12). A reproducing means 307 copies or regenerates the data on the half-failed disk drive (step S13), and stores the reproduced data on a spare disk drive 103c (step S14). The executing means 304 disconnects the half-failed disk drive 103a from the disk array system (step S15), formats the data structure of the half-failed disk drive 103a (step S16), and names the half-failed disk drive 103a as a new spare disk drive (step S17). **Sasamoto**'s reproducing means simply does not provide backup copying by a second control unit through a cache when threshold monitoring has been reached. As such, **Sasamoto** does not provide "a first control unit provided in the disk adapter for observing occurrence of access error with respect to the data disk drives, the first control unit, when the frequency of occurrence of the access error exceeds a predetermined threshold, copying data stored in the data disk drive exceeding the threshold in the spare disk drive via the cache memory" as recited in claim 1 and paraphrased as corresponding steps in claims 11 and 17.

U.S. Patent No. 6,859,888 to **Furuya** et al. describes a data storage array apparatus storing error information. FIG. 1 illustrates a data storage system including an external access apparatus 22, an input apparatus 23, and a data storage array apparatus 20 (col. 4, lines 47 – 51). In a write operation, the array apparatus 20 generates error information which indicates the existence of corrupt or invalid data and temporarily stores the error information in internal memory (col. 4, lines 61 – 67). FIG. 2 illustrates storage areas of each of the magnetic disks 1-5, which may be configured as a RAID (col. 5, lines 24 – 26). A memory 7 of the data storage array apparatus 20 stores information indicating a result of write

operations to the magnetic disks 1-5 (error information), a number of write operations from the outside, and a flag for data reconstruction (col. 5, lines 48 – 53). FIG. 4 illustrates, if the number of times of external data write operations is 3 (=threshold value) or above (Step 45), there is no external access request in the queue (Step 46), and reconstruction of data does not occur (Step 47), then the processor 6 reads the bit map data from the error table 7b and writes the bit map data to the parity group 200p across the magnetic disks 1-5 (Step 48) (col. 6, lines 44-50). At Step 48, if any write errors do not occur at the last three (=threshold value) external data write operations, a processor 6 of the data storage array apparatus 20 may not perform a bit-map-data write operation to the parity group 200p across the magnetic disks 1-5 (col. 7, lines 23-27). **Furuya**'s frequency threshold monitoring of external data write operations simply does not respond to storage disk drive access error conditions by copying data through a cache for retrieval by a spare storage disk device. As such, **Furuya** does not provide "a first control unit provided in the disk adapter for observing occurrence of access error with respect to the data disk drives, the first control unit, when the frequency of occurrence of the access error exceeds a predetermined threshold, copying data stored in the data disk drive exceeding the threshold in the spare disk drive via the cache memory" as recited in claim 1 and paraphrased as corresponding steps in claims 11 and 17.

U.S. Pat. App. Pub. No. 2002/0162057 of **Talagala** describes a storage system monitoring data integrity with each detected data integrity error stored in a count. FIG. 1 illustrates a data processing system 100 including a host 102 coupled to a storage system 106 via a host/storage connection 104 ([0018]). The storage system 106 is a hardware RAID storage subsystem or a JBOD ("Just a Bunch of Disks"). A storage device array 108 of the storage system 106 has three storage devices 110. The storage system 106 further includes an array controller 112 coupled to each of the storage devices 110 in the storage array 108 via a data path 114 ([0020]). FIG. 4 illustrates a forced failure process, wherein retrieved data is compared to redundancy data to determine data integrity ([0024]). If a demerit count COUNT has reached the storage device 110's threshold limit, then the storage device 110 is placed into a forced failure state in operation 408 ([0029]). The forced failure state can be defined as one in which general read/write commands are not directed to the drive. Spare restoration storage device 110' might be provided in storage device array 108 as a backup storage device to be used in the event that one of the other storage devices 110 in storage device array 108 fails. The reconstruction can be performed in various ways. For example, using RAID 5 techniques, parity data is calculated on an aligned group of data spread across

several hard drives. If one of the hard drives fails, a parity rebuild process is performed, in which the data from the remaining disks, including the parity data, is retrieved and an XOR calculation is utilized to rebuild all of the data from the failed hard drive. In other embodiments, all of the data is read from the force failed storage device 110 and checked using redundancy data. The data which matches with the redundancy data is copied to the restoration storage device, and the data which does not match can be restored using the redundancy data ([0031]). **Talagala**'s forced failure process simply does not respond to storage disk drive access error conditions by copying data through a cache for retrieval by a spare storage disk device. As such, **Talagala** does not provide "a first control unit provided in the disk adapter for observing occurrence of access error with respect to the data disk drives, the first control unit, when the frequency of occurrence of the access error exceeds a predetermined threshold, copying data stored in the data disk drive exceeding the threshold in the spare disk drive via the cache memory" as recited in claim 1 and paraphrased as corresponding steps in claims 11 and 17.

U.S. Pat. App. Pub. No. 2003/0233613 of **Ash** et al. describes a method and system for preventing data loss in non-volatile storage. FIG. 1 illustrates a storage subsystem 102 receiving I/O requests from hosts (104a, 104b... 104n) and connected to a storage system 106 (including disk drives 108a, 108b... 108n) ([0020]). The storage subsystem 102 has a central processing unit (CPU) 110, a cache 112 comprising a volatile memory to store tracks, and a non-volatile storage unit (NVS) 114 in which certain dirty or modified tracks in the cache 112 are buffered. Id. FIG. 4 illustrates the logic in the storage subsystem 102 that initiates an NVS scrub process 120 ([0031]). The NVS scrub process 120 checks (at block 408) the restored data (from the NVS copy of the modified copy) for errors, and then discards (at block 412) the NVS copy of the modified data. **Ash's** NVS scrub process 120 only determines whether the NVS 114 has reached a predetermined threshold of errors. If the NVS 114 has reached the predetermined threshold of errors, there is a likelihood of potential future errors in the NVS 114 that may lead to a loss of data. For example, the NVS 114 or other components such as host bus adapters connecting the storage subsystem 102 to the hosts 104a...104n could be individually replaced and the defective component isolated ([0032]). **Ash** only concerns errors occurred in a non-volatile storage unit (NVS) 114, rather than any access error with respect to the data disk drives 108. As such, **Ash** replaces or isolates the NVS 114 or host bus adapters connecting the storage subsystem 102, rather than backing-up a disk drive or generating a new parity. **Ash**'s NVS scrub process does not respond to storage

disk drive access error conditions by copying data through a cache for retrieval by a spare storage disk device. As such, **Ash** does not provide "a first control unit provided in the disk adapter for observing occurrence of access error with respect to the data disk drives, the first control unit, when the frequency of occurrence of the access error exceeds a predetermined threshold, copying data stored in the data disk drive exceeding the threshold in the spare disk drive via the cache memory," "a second control unit provided in the disk adapter for processing access request directed to the RAID group during the copying process by the first control unit, the second control unit making the backup storage take over a write request directed to the RAID group," and "a third control unit provided in the disk adapter for copying data written in the backup storage by the second control unit to the data disk drives and the spare disk drive other than the data disk drive exceeding the threshold when the copying process by the fist control unit is finished" as recited in claim 1 and paraphrased as corresponding steps in claims 11 and 17.

U.S. Pat. App. Pub. No. 2004/0153844 of **Ghose** et al. describes a failure analysis method and system for storage area networks. FIG. 1 illustrates a storage area network (SAN) 100 including switch fabrics 104, 105 ([0023]). Host servers 102 are coupled (through the fabric 104) via links 150 to individual UPEs of a storage virtualization controller 126. Storage devices 106 are coupled (through the fabric 105) via links 152 to individual downstream processing elements (DPEs) of the controller 126 ([0026]). FIG. 4 illustrates a combination of system error events, such as a fan failed event 404, an over-temperature event 406, an input-output error event 408, which then are translated error codes ([0040]). A failure analysis module 403 compares the temporal arrangement of error events against patterns in a rule 412, then executes error action associated with the rule 412 ([0041]). Errors can occur during data migration across a RAID ([0049]). **Ghose** merely provides general examples, such as "*the storage virtualization controller can be instructed to migrate data from a storage element generating error events to other more reliable areas of the storage network not experiencing the error events or failures* ([0049])", rather than as specific as responding to storage disk drive access error conditions by copying data through a cache for retrieval by a spare storage disk device, etc. As such, **Ghose** does not provide "a first control unit provided in the disk adapter for observing occurrence of access error with respect to the data disk drives, the first control unit, when the frequency of occurrence of the access error exceeds a predetermined threshold, copying data stored in the data disk drive exceeding the threshold in the spare disk drive via the cache memory," "a second control unit provided in

the disk adapter for processing access request directed to the RAID group during the copying process by the first control unit, the second control unit making the backup storage take over a write request directed to the RAID group," and "a third control unit provided in the disk adapter for copying data written in the backup storage by the second control unit to the data disk drives and the spare disk drive other than the data disk drive exceeding the threshold when the copying process by the fist control unit is finished" as recited in claim 1 and paraphrased as corresponding steps in claims 11 and 17.

Conclusion

Based on the results of the comprehensive prior art search as discussed above, Applicants contend that the disk array system as now recited in independent claims 1, especially the features of "a first control unit provided in the disk adapter for observing occurrence of access error with respect to the data disk drives, the first control unit, when the frequency of occurrence of the access error exceeds a predetermined threshold, copying data stored in the data disk drive exceeding the threshold in the spare disk drive via the cache memory," "a second control unit provided in the disk adapter for processing access request directed to the RAID group during the copying process by the first control unit, the second control unit making the backup storage take over a write request directed to the RAID group," and "a third control unit provided in the disk adapter for copying data written in the backup storage by the second control unit to the data disk drives and the spare disk drive other than the data disk drive exceeding the threshold when the copying process by the fist control unit is finished" are patentably distinct from the cited prior art references.

In particular, as now recited in the claim 1 (for example, the embodiment shown in Figs. 1 & 5), the disk array system 10 comprises: a channel adapter 11 for controlling data transfer with respect to a host device 1; a plurality of data disk drives 16 configuring a RAID group 17; at least one spare disk drive 16(SP) provided as a spare for the data disk drives 16; a disk adapter 12 for controlling data transfer with respect to the data disk drives 16 and the spare disk drive 16(SP); a cache memory 14 used by the channel adapter 11 and the disk adapter 12 for storing data; a control memory 13 used by the channel adapter 11 and the disk adapter 12 for storing control information; a backup storage 17(S) provided separately from the data disk drives 16 and the spare disk drive 16(SP); a first control unit provided in the disk adapter 12 for observing occurrence of access error with respect to the data disk drives 16, the first control unit, when the frequency of occurrence of the access error exceeds a

predetermined threshold, copying data stored in the data disk drive 16 exceeding the threshold in the spare disk drive 16(SP) via the cache memory 14; a second control unit provided in the disk adapter 12 for processing access request directed to the RAID group 17 during the copying process by the first control unit, the second control unit making the backup storage 17(S) take over a write request directed to the RAID group 17; and a third control unit provided in the disk adapter 12 for copying data written in the backup storage 17(S) by the second control unit to the data disk drives and the spare disk drive other than the data disk drive exceeding the threshold when the copying process by the fist control unit is finished.

Claim 1 is representative and is directed toward a disk array system including first, second, and third control units, with the first control unit observing frequency of disk access error above a threshold and copying data into cache memory in response, the second control unit processing access requests and making backup storage during the copying, and a third control unit copying data in the backup storage. Claims 11 and 17 are directed to the same invention and set forth operations as methods of steps. With reference to the claims, a disk array system or method comprises a first control unit (or step) observing a frequency of disk access error above a threshold and copying data into cache memory in response, a second control unit (or step) processing access requests and making backup storage during the copying, and a third control unit (or step) copying data form the backup storage.

In view of all the above, clear and distinct differences as discussed exist between the present invention as now claimed and the prior art references, Applicants respectfully contend that the prior art references cannot anticipate the present invention or render the present invention obvious. Rather, the present invention as a whole is distinguishable, and thereby allowable over the prior art.

Favorable consideration of this application is respectfully solicited. Should there be any outstanding issues requiring discussion that would further the prosecution and allowance

of the above-captioned application, the Examiner is invited to contact the Applicants' undersigned representative at the address and phone number indicated below.

Respectfully submitted,

Stanley P. Fisher
Registration Number 24,344

Juan Carlos A. Marquez
Registration Number 34,072

REED SMITH LLP
3110 Fairview Park Drive
Suite 1400
Falls Church, Virginia 22042
(703) 641-4200

April 18, 2005

SPF/JCM/JT